



Field Testing the User Interface Standard

The purpose of this discussion is to present our proposal to the Professional Advisory Committee (PAC) for field testing the interface standard. PAC members can provide comments at a telephone meeting to be scheduled for early June, or to Bruce before that meeting. At the meeting we expect to amend this as needed then gain PAC approval for it. You should keep in mind that there are limits on LBNL's resources – and even greater constraints on the calendar time – that can be devoted to this testing. If other organizations can do testing, that is helpful for the standard, but would need to happen soon to fit into our project schedule which calls for the standard content to be fixed by early October at the latest. We also present information about test methods prescribed for international standard symbols.


This plan follows from the testing memo of April 9 and the discussion on the April 10 PAC meeting. See the publications page of the project web site for the April 9 memo and the meeting summary.

The testing done at UC Berkeley¹ covered many parts of the standard at the same time. This made both asking questions of subjects and discussing the results confusing. For this next phase of testing it is clear that we should only test parts of the standard individually, and assume as given the rest of the standard. This greatly simplifies the questions and interpretation. For example we can tell people that devices are on, off, or in a sleep mode when asking about indicator light colors rather than having the terminology for low-power modes and the color to be associated both up in the air.

There are six key principles of the standard. The PAC identified two principles that merit testing:

- Drop the  symbol from use and redefine  to mean “power”.
- Adopt the “green/amber/off” color indications for power indicators.

There was not interest in or apparent need for testing the other four principles:

- Use only three power states when possible: On, Off, and Sleep.
- Use the word “Power” for terminology about power.
- Use the “sleep” metaphor for entering, being in, and coming out of low-power states; use the moon symbol –  – to mean sleep.
- Clarify that “hibernate” is off; a new term is needed.

For hibernate=off, the term to use for hibernate does remain in question. However, this testing exercise is not an appropriate venue to choose a new term. In this project we have been identifying existing designs that deserve to be made the basis for the voluntary standard. Identifying a new term for hibernate would be more akin to designing a product, which we don't see appropriate for this process.

For each of the two tests we propose the following 4-step process:

1. Summarize existing usages as they appear on different product types.
2. Pre-testing: Experiment with different ways of asking people about indicator light options to identify some promising methods.
3. Local testing: Survey a sample of convenience (several dozen people in and around Berkeley, CA) to get initial results and identify the most useful questions and form of questions to ask.
4. Web-based testing: Create a survey on web pages of a limited set of questions that best help inform our decision about the standard. This would probably be made up entirely multiple choice responses.

¹ See the publications page for a summary and link to the full studies.

For the web-based testing we would report on the entire pool of subjects as well as by a limited number of subject characteristics such as country of residence. The survey instrument would be like that used in the UCB study, but less complex.

For the web study, we would like to cast a wide net in soliciting subjects, particularly with the goal of getting non-U.S. subjects. We would also like to make the survey instrument available in languages other than English. For both of these, it would be necessary to get PAC member companies (or others) to do the translations and solicitation.


After each stage of the process, we would reconsider what further stages made sense, modifying them or terminating them as appropriate.


Test A: Power Symbols — and



Goal

The goal is to help understand the effectiveness and value of maintaining the two symbols with their current meanings.

Discussion

The UCB testing found that people seem most comfortable pressing the  button (for any purpose!) and that they will readily press a power button to wake up a PC².

The existing symbols were designed with switches in mind, but manufacturers now used routinely use them for both switches, indicator lights, and—at least for —a power state. These multiple usages pose challenges for definition and use.

The existing international symbol standards are not well defined, but have to be interpreted by those who are concerned with safety indication. These people usually interpret the definitions them to mean that  is for an on/off switch that goes to zero watts, and  is for one that goes to something other than zero. For this to be effective for consumer products, people must:

- notice the difference,
- understand the difference in meaning,
- care enough to think about it, and
- be doing things (e.g. taking the cover off the device to expose electronics) that make it relevant.

Procedure

For initial testing, we propose to:

1. Present subjects with one of the symbols, then ask if they recognize it and what they think it means.
2. Repeat the process with the other symbol (varying the order).
3. Explain that one is to mean a “power button”/switch for devices which go to zero energy in off and one for those that still draw a small amount of power when off.
4. Ask if they know which is which.
5. Ask if they care or not, and why.
6. Ask which they think is most common for a power button and finally, which they prefer for a power button.

² We are assuming that the moon is to be the symbol for sleep and that neither of the power symbols are to be used for sleep, so sleep and the moon need not be part of this test.

Test B: Power Indicators

Goal

The goal is to identify associations with individual colors, and with certain colors blinking or not. This is to help clarify which indication schemes are viable and which are preferable.

Discussion

The UCB study found that for laptops, more people would poke or prod it than check indicator lights or the display. For copiers, many more checked the display or indicators. Indicator lights alone were relied on by only a sixth and a third of people respectively. That indicators did not rate higher might be due to their current inconsistency.

Subjects were asked about the correspondence between device state and indicator color. Questions were asked both ways — what color corresponds to a particular state, and what state corresponds to each color. They were first asked what colors were associated with each state or condition.

- Green was overwhelmingly the choice for on, and “no light” was the dominant choice for off. For off, red made a respectable showing, though that may have been due to subjects reading the question as addressing the color of an on/off button (like a STOP sign) rather than the color of an indicator light.
- For sleep there was the least clarity, with blue rating higher than yellow (possibly this is due to yellow not standing out on the white background of the survey). This is odd as we have never seen blue as a sleep indicator on any device.

Subjects were asked about associations with green, orange, yellow, and red, and blinking versions of each. Why the students chose to ask about both orange and yellow isn't clear; the fact that both were present may have led some subjects to think that yellow should mean something different from orange.

- Green was overwhelmingly identified with on.
- Interestingly, non-blinking red was never identified with Attention/Input or Error, but rated the highest on Don't Know.
- Non-blinking was rarely identified with Attention/Input, Error, or Transitions, which is good evidence for blinking for these indications.
- Aside from red, blinking rated well for sleep, but not as high as constant orange or yellow.

From all this we think it is reasonable to tell people that a solid green light is on and off is off. Since red is to be reserved for error conditions or faults by many standards, we see no need to include it as a color for a power indicator in these tests (and suggests that red for power indicators should be phased out in consumer electronics). The CIE standards for traffic signal lights include specifications for red, green, yellow, blue, and white. Eliminating red and white we are left with green, yellow, and blue. We would assure that colors we use in testing are consistent with recommendations that maximize the ability of color-deficient people to recognize the difference between colors, particularly green and yellow (this is a problem with some current products).

For blinking, there are four common variants that could be included:

- Equal length on and off cycles of about 1 second total.
- Short on cycles (e.g. 10% of total) with the total about 2 seconds.
- “Breathing” or “pulsating” (as on recent Apple computers)
- Fast blinking (e.g. 5 Hz).

Procedure

Clarify up front that we are talking about “Power Indicator Lights”, usually near a “Power Button” on office equipment or consumer electronics. We will further state that devices are on, asleep, or off.

First, we will ask for associations. We will probably begin with open-ended questions, but eventually arrive at a fixed set for people to choose from. The UCB study used On, Off, Sleep, Attention/Input, Error, Don’t Know, and Transition State.

- First, green (to get them warmed up)
- Then, blue (to clear their mind)
- Then, flashing green or steady yellow (vary order randomly), then the other
- Then, flashing yellow
- Then, breathing green.

Second, we tell people that devices are On, Asleep, or Off, and for each option ask if it is clear, and if they like it. Options we know will be used are green/flashing green/off, and green/yellow/off.

Test methods for ISO/IEC standards

ISO 9186, “Graphical Symbols — Test methods for judged comprehensibility and for comprehension” (Second edition, April 2001) specifies procedures to be used in advance of establishing international standard symbols. It is *not* our intention to conduct our tests for this purpose, but we may find it helpful to align our tests with this standard in part.

Tests can be of two types: “comprehension judgment”, what percent of others will understand a symbols, and “comprehension”, whether the subject herself/himself understands it correctly. Our tests will be a mixture of the two. The “referent” of a symbol is the “idea or object that the graphical symbol is intended to represent”.

Depending on the test, testing in two or three countries is required, with the goal that they be of substantially different cultures.

The standard provides for computer display based testing.

Symbols are to be presented in random orders.

At least 50 “respondents” are required for each country a test is done in.

Respondents are to be:

- Representative of the user population.
- Presented with both two test types about the same symbol or referent.
- Told the context in which the symbol will usually be used in, then asked what they think it means, and (if applicable) what action they should take.
- Asked for their age (by a broad range), gender, education, cultural background, and where relevant, physical ability.